Chapter 13

Quality Assessment and Statistical Analysis of Air Monitoring Data

Chapter 13
Quality Assessment and Statistical Analysis of Air Monitoring Data
Table of Contents

1.0 Introduction

The Indiana Department of Environmental Management, Office of Air Quality, Quality Assurance Section (IDEM, OAQ, QAS), must submit to the United States Environmental Protection Agency (U.S. EPA), Region V, within 90 days after the end of each quarter, the results of all required valid precision, bias, and accuracy tests conducted within the state during that quarter. In order to meet the 90-day deadline, all precision, bias, and accuracy data from each reporting agency must be submitted to the OAQ/QAS within 30 days after the end of each quarter.

Ambient data must be submitted to the IDEM Ambient Monitoring Section within 60 days after the end of the quarter. This will allow sufficient time for review before the AQS submittal deadline of 90 days after the end of each reporting calendar quarter.

All ambient and precision and accuracy data submittals are done electronically. Additional information and other statistical calculations not mentioned in this chapter may be found in other chapters of this manual and in U.S. EPA Guidance Documents.

NOTE: U.S. EPA performs some calculations when the ambient and audit data is submitted.

In order to achieve national uniformity in data reliability for reporting organizations, U.S. EPA has set upper and lower 95-percent probability limits by parameter. These probability limits are listed in Table 1. Additional flagging criteria can be found in the individual sections.

Table 1
Probability Limits

Parameter	Precision Limit	Accuracy Limit
PM_{10} (Manual)	± 15%*	± 20%
PM_{10} (Automated)	± 15%*	± 20%
$PM_{2.5}$ (Manual)	See Individual Sections	See Individual Sections
Pb	± 15%	± 20%
SO_2	± 15%	± 20%
CO	± 15%	± 20%
NO_2	± 15%	± 20%
O_3	± 15%	± 20%
PAMS VOC	See Individual Sections	See Individual Sections

Data falling outside these limits are unreliable. Use of such data for State Implementation Plans (SIPs) or as a basis for regional classifications would be highly questionable by the permitting authority. Whenever the 95 percent probability limits begin to approach or exceed the limits discussed, determine the source(s) of the problem and implement corrective action. In addition, at low concentrations agreement between collocated samplers may be relatively poor when expressed as percent differences. For this reason, collocated measurement pairs are selected for use in the precision calculations only when both measurements are above the limits noted in Table 2.

Table 2
Above Probability Limits

Parameter	Limit
PM_{10}	20 µg/m^3
$PM_{2.5}$	6 µg/m^3
Pb	0.15 µg/m^3

The QAS recommends all reporting agencies calculate parameter precision, bias, accuracy, and perform other statistical analysis on a continuing basis. Calculate upgraded statistics after each audit for each analyzer/pollutant, and for PM_{10}, $PM_{2.5}$, and Pb, as soon as analytical results become available. The following sections explain the methods used to perform statistical analysis (e.g. precision, bias, and accuracy) for automated and manual methods of monitoring ambient air quality. ***Formulas for all Equations referenced in the following sections are contained in Appendix A.***

2.0 Automated Methods for Gas Pollutants

2.1 Precision

Estimates of data precision for automated methods of SO_2, NO_2, CO, and O_3 are calculated from the results of biweekly audits (audits performed every two weeks). These audits must include a precision checkpoint between 0.08 - 0.10 ppm for SO_2, NO_2, and O_3. The precision point for CO must be between 8-10 ppm. Audit procedures are listed by parameter in other chapters of this manual.

2.1.1 Single Instrument Precision

Calculate the percent difference (d_i) for each precision check using Equation 1, where Y_i is the concentration indicated by the analyzer for the i-th precision check and X_i is the known concentration for the i-th precision check.

For each instrument j, calculate the running average (d_j) and standard deviation (S_j) using Equations 2 and 3. The number of precision checks to date is represented by n.

2.1.2 Precision for Reporting Organization

For each gas pollutant, calculate the average of the averages (D) and the pooled standard deviation (S_a), using either Equations 4 and 5 or 4a and 5a.

Note: Use Equations 4 and 5 when the same numbers of precision checks are made for each instrument. When instruments have an unequal number of audits, use Equations 4a and 5a to obtain a weighted average and standard deviation. The number of analyzers audited within the reporting organization is represented by *k*.

For each gas pollutant, the 95% probability limits for the precision for the reporting organization are calculated using Equations 6 and 7.

2.2 Accuracy

Estimates of the accuracy of automated methods are calculated from the results of independent audits (See Chapter 1 for accuracy audit procedures). Although not required by the U.S. EPA, the QAS recommends that all reporting agencies perform accuracy audits not less than once each quarter for each analyzer (See chapters 1 and 14 for Prevention of Significant Deterioration site requirements). More frequent accuracy audits are encouraged to help troubleshoot any problems and to see how linear the analyzer is over the whole sampling range. The audit device used for the accuracy audit must be different from the device used to calibrate the analyzer.

2.2.1 Single Analyzer Accuracy

Calculate the percent difference (d_i) for each audit concentration level using Equation 1.

Note: The correct concentration levels for accuracy audits are listed below in Table 3.

Table 3
Concentration Levels for Accuracy Audits

Parameter	Concentration	Level
CO	3 - 8 ppm	1
CO	15 - 20 ppm	2
CO	35 - 45 ppm	3
CO	80 - 90 ppm	4
SO_2, O_3, NO_2	0.03 - 0.08 ppm	1
SO_2, O_3, NO_2	0.15 - 0.20 ppm	2
SO_2, O_3, NO_2	0.35 - 0.45 ppm	3
SO_2, O_3, NO_2	0.80 - 0.90 ppm	4

2.2.2 Accuracy for Reporting Organization

Using Equation 8, calculate for each audit concentration level the average (D) of the individual percent differences (d_i) for all analyzers (n) measuring a given pollutant.

Using Equation 9, calculate for each audit concentration level the standard deviation (S_a) of all the different percent differences for all analyzers audited. Repeat this for each pollutant.

For each pollutant, calculate the 95 percent probability limits for the accuracy of the reporting organization at each concentration level using Equations 6 and 7.

3.0 Manual and Automated Methods for PM_{10}

3.1 Precision

Estimates of precision for manual methods of monitoring ambient air quality for PM_{10} are calculated from the results obtained from the collocation of two samplers. On an on-going basis, calculate a combined precision probability interval so that corrective action can be taken before precision limits are reached.

Because of the high cost of providing a collocated PM_{10} automated sampler, flow checks are used instead to assess precision. A one-point check of each PM_{10} samplers normal operating flow rate is performed. If a precision is made in conjunction with a zero or span adjustment, it must be made prior to such zero and span adjustment.

3.1.1 Single Sampler Precision

For each selected measurement pair, the percent difference (d_i) is calculated using Equation 10. The concentration of the collocated sampler is represented by y_i, and the concentration of the primary sampler designated for reporting air quality for the site is represented by x_i.

Using Equations 2 and 3, calculate the average percent differences (d_j) and the standard deviation (S_j) where n = the number of selected measurement pairs at each site.

3.1.2 Precision for Reporting Organization

Using Equations 4 and 5 or 4a and 5a, calculate the average percent differences (D) and pooled standard deviation (S_a). The variable k in the aforementioned Equations is the number of collocated sites. Calculate the 95 percent probability limits using Equations 11 and 12.

3.2 Accuracy

Estimates of the accuracy of manual and automated methods for PM_{10} are calculated from the results of independent audits (See Chapter 7 for audit procedures). All devices used in the accuracy audit must be different from those used during the calibration of samplers and analysis of samples.

3.2.1 Single Sample Accuracy

For the flow rate audit described in chapter 7, the percent difference (d_i) for each audit is calculated using Equation 1 where x_i = the known flow rate and y_i = the flow rate indicated by the sampler.

3.2.2 Accuracy for Reporting Organization

For each type of particulate matter sampled, the average (D) of the individual percent differences for all similar samplers is calculated using Equation 8. The standard deviation (S_a) of the percent differences for all of the similar particulate samplers is calculated using Equation 9, and the 95 percent probability limits are calculated using Equations 6 and 7.

4.0 Precision, Accuracy, and Bias for Manual $PM_{2.5}$

4.1 Quality Control Checks and Audits

The only data that reporting organizations are required to report is the network collocated data. The U.S. EPA will report the collocated **Federal Reference Method (FRM)** data. The U.S. EPA uses the reported results to derive precision, accuracy, and bias estimates according to the following procedures listed below.

4.2 Collocated Samplers

Collocated sampler results are used to assess measurement system precision. A collocated sampler pair consists of a primary sampler (used for routine monitoring) and a duplicate sampler (used as a quality control check). Quarterly precision estimates are calculated for each pair of collocated samplers and for each method designation employed by each reporting organization. Annual precision estimates are calculated for each primary sampler, for each EPA Federal reference method and equivalent method designation employed by each reporting organization, and nationally for each U.S. EPA Federal reference method and equivalent method designation. Collocated measurement pairs are selected for use in the precision calculations only when both measurements are above 6 $\mu g/m^3$.

4.3 Percent Difference for a Single Check (d_i)

The percentage difference, d_i, for each check is calculated using Equation 13, where x_i represents the concentration produced from the primary sampler, and y_i represents concentration reported for the duplicate sampler.

4.4 Coefficient of Variation (CV) for a Single Check (CV_i)

The Coefficient of variation, CV_i, for each check is calculated by dividing the absolute value of the percentage difference, d_i, by the square root of two as shown in Equation 14.

4.5 Precision of a Single Sampler - Quarterly Basis ($CV_{j,q}$)

For particulate sampler *j*, the individual coefficients of variation ($CV_{j,q}$) during the quarter *q* are pooled using Equation 15, where $n_{j,q}$ is the number of pairs of measurements from collocated samplers during the quarter.

The 90 percent confidence limits for the single sampler's CV are calculated using Equations 16 and 17, where $\chi^2_{.05,df}$ and. $\chi^2_{.95,df}$ are the 0.05 and 0.95 quantiles of the chi-square (χ^2) distribution with $n_{j,q}$ degrees of freedom.

4.6 Precision of a Single Sampler - Annual Basis

For particulate sampler *j*, the individual coefficients of variation, CV_i, produced during the calendar year are pooled using Equation 15, where n_j is the number of checks made during the calendar year. The 90 percent confidence limits for the single sampler's CV are calculated using Equations 16 and 17, where $\chi^2_{.05,df}$ and $\chi^2_{.95,df}$ are the 0.05 and 0.95 quantiles of the chi-square (χ^2) distribution with $n_{j,q}$ degrees of freedom.

4.7 Precision for each EPA Federal Reference Method and Equivalent Method Designation Employed by each Reporting Organization - Quarterly Basis ($CV_{k,q}$)

For each method designation *k* used by the reporting organization, the quarter's single sampler coefficients of variation, $CV_{j,q}$s obtained with Equation 15, are pooled using Equation 18, where $n_{k,q}$ is the number of collocated primary monitors for the designated method (but not collocated with FRM samplers) and $n_{j,q}$ is the number of degrees of freedom associated with $CV_{j,q}$.

The number of method CVs produced for a reporting organization will equal the number of different method designations having more than one primary monitor employed by the organization during the quarter. (When exactly one monitor of a specified designation is used by a reporting organization, it will be collocated with an FRM sampler.)

4.8 Precision for each Method Designation Employed by each Reporting Organization - Annual Basis (CV_k)

For each method designation k used by the reporting organization, the quarterly estimated coefficients of variation, $CV_{k,q}$, are pooled using Equation 19, where $n_{k,q}$ is the number of collocated primary monitors for the designated method during the qth quarter and also the number of degrees of freedom associated with the quarter's precision estimate for the method designation $CV_{k,q}$.

4.9 FRM Audits

FRM audits are performed once per quarter for selected samplers. The reporting organization reports concentration data from the primary sampler. Calculations for FRM audits are similar to those for collocated samplers having FRM samplers as duplicates. The calculations differ because only one check is performed per quarter.

4.10 Sequential Samplers

For sequential samplers, the sampler shall be configured for the maximum number of sequential samples and shall be set for automatic collection of all samples sequentially such that the test samples are collected equally, to the extent possible, among all available sequential channels or utilizing the full available sequential capability.

5.0 Test Procedures for Methods for Lead

5.1 Sample Collection

Collect simultaneous 24-hour samples (filters) of lead at the test site or sites using proper reference methods until at least 10 filter pairs have been obtained. If the conditions of 40 CFR Part 53.30 (d) (4) apply, collect at least 10 common samples (filters) in accordance with 40 CFR Part 53.30 (d) (4) and divide each to form the filter pairs.

5.2 Audit Samples

Three audit samples must be obtained from the address given in 40 CFR Part 53.4 (a). The audit samples are 3/4 x 8 inch glass fiber strips containing known amounts of lead at the following nominal levels: 100 µg/strip; 300 µg/strip; 750 µg/strip. The true amount of lead, in total µg/strip, will be provided with each audit sample.

5.3 Filter Analysis

For each audit sample, analyze each filter extract three times in accordance with the reference method analytical procedure. The analysis of replicates should not be performed sequentially, i.e., a single sample should not be analyzed three times in sequence. Calculate the indicated lead concentrations for the reference method samples in µg/m^3 for each analysis of each filter. Calculate the indicated total lead amount for the audit samples in µg/strip for each analysis of each strip. Label these test results as R_{1A}, R_{1B}, R_{1C}, R_{2A}, R_{2B}, ..., Q_{1A}, Q_{1B}, Q_{1C}, ..., where R denotes results from the reference method samples; Q denotes results from the audit samples; 1, 2, 3 indicate the filter number, and A, B, C indicate the first, second, and third analysis of each filter, respectively.

5.4 Average Lead Concentration

Calculate average lead concentrations for each filter by averaging the concentrations calculated from the three analysis using Equation 20, where *i* is the filter number.

5.5 Test for Precision

Calculate the precision (P) of the analysis (in percent) for each filter, as the maximum minus the minimum divided by the average of the three concentration values using Equation 21 or Equation 22, where *i* indicates the filter number.

If precision value (P_{Ri}) exceeds 15 percent, the precision of the reference method analytical procedure is out of control. Corrective action must be taken to determine the source(s) of imprecision and the reference method determinations must be repeated according to Section 5.3 of this chapter, or the entire test procedure (starting with Section 5.1 of this chapter) must be repeated.

5.6 Test for Accuracy

For the audit samples, calculate the average lead concentration for each strip by averaging the concentrations calculated from the three analyses using Equation 23, where *i* is the audit sample number.

Calculate the percent difference (D_q) between the indicated lead concentration for each audit sample and the true lead concentration (T_q) using Equation 24.

If any difference value (D_{qi}) exceeds ±5 percent, the accuracy of the reference method analytical procedure is out of control. Corrective action must be taken to determine the source of the error(s) (e.g., calibration standard discrepancies, extraction problems, etc.) and the reference method and audit sample determinations must be repeated according to Section 5.3, or the entire test procedure (starting with Section 5.1) must be repeated.

5.7 Flow Accuracy

In addition to the precision and accuracy of the lead analysis, the flow rate on the lead sampler will have an accuracy audit performed on it. This procedure is the same as described for the PM_{10} sampler in Section 3.2 of this chapter.

6.0 PAMS Volatile Organic Compounds (VOC)

Consult Chapter 8 of the Indiana Department of Environmental Management, Office of Air Quality, Quality Assurance Manual and the Environmental Protection Agency Technical Assistance Document for details on precision and accuracy requirements.

7.0 Submittal of Precision and Accuracy Data

The precision and accuracy reports should be submitted to the Indiana Department of Environmental Management, Office of Air Quality, Quality Assurance Section. This information will be reviewed and submitted to EPA's AQS system.

APPENDIX A - EQUATIONS

Equation 1:

$$d_i = \frac{Y_i - X_i}{X_i} \times 100$$

Equation 2:

$$d_j = \frac{1}{n} \sum_{i=1}^{n} d_i$$

Equation 3:

$$S_j = \sqrt{\frac{1}{n-1}\left[\sum_{i=1}^{n} d_i^2 - \frac{1}{n}\left[\left(\sum_{i=1}^{n} d_i\right)^2\right]\right]}$$

Equation 4:

$$D = \frac{1}{k}\sum_{j=1}^{k} d_j$$

Equation 4a:

$$D = \frac{n_1 d_1 + n_2 d_2 + \cdots + n_j d_j + \cdots + n_k d_k}{n_1 + n_2 + \cdots + n_j + \cdots + n_k}$$

Equation 5:

$$S_a = \sqrt{\frac{1}{k}\sum_{j=1}^{k} S_j^2}$$

Equation 5a:

$$S_a = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2 + \cdots + (n_j - 1)S_j^2 + \cdots (n_k - 1)S_k^2}{n_1 + n_2 + \cdots + n_j + \cdots n_k - k}}$$

Equation 6:

$$\textit{Upper 95\% probability limit} = D + 1.96 S_a$$

Equation 7:

$$\textit{Lower 95\% probability limit} = D - 1.96 S_a$$

Equation 8:

$$D = \frac{1}{n}\sum_{i=1}^{n} d_i$$

Equation 9:

$$S_a = \sqrt{\frac{n\sum_{i=1}^{n} d_i^2 - \left(\sum_{i=1}^{n} d_i\right)^2}{n(n-1)}}$$

Equation 10:

$$d_i = \frac{y_i - x_i}{\frac{1}{2}\left(y_i + x_i\right)}$$

Equation 11:

$$\textit{Upper 95\% probability limit} = D + 1.96 S_a$$

Equation 12:

$$\textit{Lower 95\% probability limit} = D - 1.96 S_a$$

Equation 13:

$$d_i = \frac{y_i - x_i}{\frac{1}{2}\left(y_i + x_i\right)}$$

Equation 14

$$CV_i = \frac{|d_i|}{\sqrt{2}}$$

Equation 15:

$$CV_{j,q} = \sqrt{\frac{\sum_{i=1}^{n_j} CV_i}{n_{j,q}}}$$

Equation 16:

$$\textit{Upper Confidence Limit} = CV_{j,q}\sqrt{\frac{n_{j,q}}{\chi^2_{0.05,n_{j,q}}}}$$

Equation 17:

$$\textit{Lower Confidence Limit} = CV_{j,q}\sqrt{\frac{n_{j,q}}{\chi^2_{0.95,n_{j,q}}}}$$

Equation 18:

$$CV_{k,q} = \sqrt{\frac{\sum_{j=1}^{n_{k,q}} (CV_{j,q}^{2} \times n_{j,q})}{\sum_{j=1}^{n_{j,q}} n_{j,q}}}$$

Equation 19:

$$CV_{k} = \sqrt{\frac{\sum_{q=1}^{4} (CV_{k,q}^{2} \times n_{k,q})}{\sum_{q=1}^{4} n_{k,q}}}$$

Equation 20:

$$R_{i\_avg} = \frac{R_{iA} + R_{iB} + R_{iC}}{3}$$

Equation 21:

$$P_{R_i} = \frac{R_{i\_\max} - R_{i\_\min}}{R_{i\_avg}} \times 100\%$$

Equation 22:

$$P_{C_i} = \frac{C_{i\_\max} - C_{i\_\min}}{C_{i\_avg}} \times 100\%$$

Equation 23:

$$Q_{i\_avg} = \frac{Q_{iA} + Q_{iB} + Q_{iC}}{3}$$

Equation 24:

$$D_{q_i} = \frac{Q_{i\_avg} - T_{qi}}{T_{qi}} \times 100\%$$

Equation 25:

$$D_{i,n} = \frac{C_{ij} - R_{ik}}{R_{ik}} \times 100\%$$